WPS4473

POLICY RESEARCH WORKING PAPER 4473

Patterns of Export Diversification in Developing Countries:

Intensive and Extensive Margins

Alberto Amurgo-Pacheco
Martha Denisse Pierola

The World Bank
International Trade Department
January 2008

Policy Research Working Paper 4473

Abstract

This paper uses highly disaggregated trade data to investigate geographic and product diversification patterns across a group of developing nations for the period from 1990 to 2005. The econometric investigation shows that the gravity equation fits the observed differences in diversification across nations. The analysis shows that exports at the intensive margin account for the most important share of overall trade growth. At the extensive margin, geographic diversification is more important than product diversification, especially for developing countries. Taking part in free trade agreements, thereby reducing trade costs, and trading with countries in the North are also found to have positive impacts on export diversification for developing countries.

This paper—a product of the International Trade Department—is part of a larger effort in the department to stimulate analysis of policies designed to increase export diversification. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at mpierola@worldbank.org, pierola3@hei.unige.ch and amurgop1@hei.unige.ch.

Patterns of Export Diversification in Developing Countries: Intensive and Extensive Margins

Alberto AMURGO-PACHECO[1] *and Martha Denisse PIEROLA*[2 *]

JEL. **Keywords**: export diversification, intensive margin, extensive margin.

JEL Codes: F13 F14 F15

Acknowledgements: We are very grateful to Richard Baldwin for his valuable contributions and to Richard Newfarmer and Paul Brenton for their comments on a previous draft. We are also indebted to Mombert Hoppe, Heiko Hesse, Roberta Piermartini and Rosen Marinov for helpful discussions.

---

[1] Graduate Institute of International Studies, Avenue de la Paix 11A, 1202 Geneva, Switzerland, email: amurgop1@hei.unige.ch
[2] Graduate Institute of International Studies, Avenue de la Paix 11A, 1202 Geneva, Switzerland, email: pierola3@hei.unige.ch

1. INTRODUCTION

Most researchers would agree that export diversification matters and it is especially important for developing countries. The literature has, so far, focused on identifying transmission channels. For example, developing countries' exports tend to be concentrated on a few products, often commodities, with very volatile demand. This translates into high income instability, which in turn provokes high growth volatility. Export diversification in this setting has the advantage of creating a more stable income inflow. Others view the benefits from diversification in terms of the spillovers in the economy as a result of having a more diversified production structure, Hausmann and Klinger (2006).

In our view, in order to be able to understand export diversification we need first to understand its microeconomic determinants. We are still lacking the necessary understanding of these determinants but recent theoretical and empirical contributions to the trade literature offer new tools to investigate this issue. In this regard Melitz (2003) develops a theoretical framework that provides a simple explanation for the presence of many of the zeros in the trade matrix. Empirical work has already been done on these grounds. For instance, Baldwin and Di Nino (2006) studied how the implementation of the Euro has affected the extensive margin of trade (the number of varieties traded) in European countries. In another recent paper, Amurgo-Pacheco (2006) focuses on the Euro-Mediterranean FTA to investigate how preferential trade agreements affect the number of varieties traded among its members (through the reduction in trade costs).

Following this line of work, the objective of the paper is to investigate the differences in export diversification patterns between developed and developing countries at a highly disaggregated level (we use the HS 6-digit trade data). For instance, a common observation is that the degree of diversification varies depending on destination-specific factors such as trade costs and the size of the market. In particular, we observe that trading with larger markets implies a higher probability of finding demand for more and new products, and thus, exporting more product varieties. We first disentangle some of the determinants of trade growth and investigate the relevance of diversification. Then, if diversification accounts for a non trivial share of export growth, we want to know what determines it and what can be done from a policy perspective. We address these points throughout the paper.

We organized the paper as follows: Section 2 introduces the dataset and discusses some stylized facts on export diversification and the margins of trade for both, the developed and developing nations. In Section 3, we present the theoretical framework used to investigate export diversification patterns. In Section 4 we present our formal empirical analysis using gravity equations to estimate trade flows. We conclude in Section 5 summarizing the main findings of the paper, discussing some policy recommendations that can be drawn from them, and providing some avenues for future research.

2. STYLIZED FACTS

This section presents stylized facts on the evolution of export diversification which help us to disentangle the relative importance of the intensive and extensive margins of trade. We begin with a description of the dataset used for both, the stylized facts, and the econometric analysis.

2.1. *The dataset*

The ideal dataset to deal with export diversification issues from a microeconomic point of view would consist of firm-level export data for each country pair; however, such data are not available to researchers. We use instead the highest level of internationally comparable disaggregated country-level trade data that are currently available, namely the 6-digit level of the Harmonised System (HS) from Comtrade database.[3] Our time frame is the period from 1990 to 2005, and the sample includes a selection of 24 developed and developing economies. The set of countries encompasses Argentina (ARG), Australia (AUS), Bangladesh (BGD), Brazil (BRA), Canada (CAN), Chile (CHL), China (CHN), EU (EEC15), Ghana (GHA), Indonesia (IDN), India (IND), Japan (JPN), Kenya (KEN), Morocco (MAR), Mexico (MEX), Mauritius (MUS), Malaysia (MYS), Peru (PER), Thailand (THA), Tunisia (TUN), Turkey (TUR), Uganda (UGA), USA and South Africa (ZAF).

The group of countries that has been selected accounts for a very significant share of world trade. In year 1995, the selected group of countries accounted for two-thirds of world trade. When looking at trade among developing countries alone, the developing countries included in the sample also account for two-thirds of the total amount of trade among developing nations.

2.2. *How do we measure diversification?*

Export diversification has been measured in many ways. For instance, it has been measured using concentration indexes, counts of exported products or even indexes that take into account the productivity content of the export basket.[4] We use some of these measures in the first part of this section, but most of our analysis rests on a much more straightforward way to look at diversification. We follow Amurgo-Pacheco (2006), and Baldwin and Di Nino (2006) to look at the intensive versus extensive composition of trade.

With respect to this latter, we define what is meant by intensive and extensive margin. The *intensive margin* of trade refers to the growth of exports in goods that are already being exported. We can refer to these as "old products". The *extensive margin* is defined as the growth of exports in new categories. In a similar way we can think of it as "new products". For the purpose of this analysis, we have defined old products as

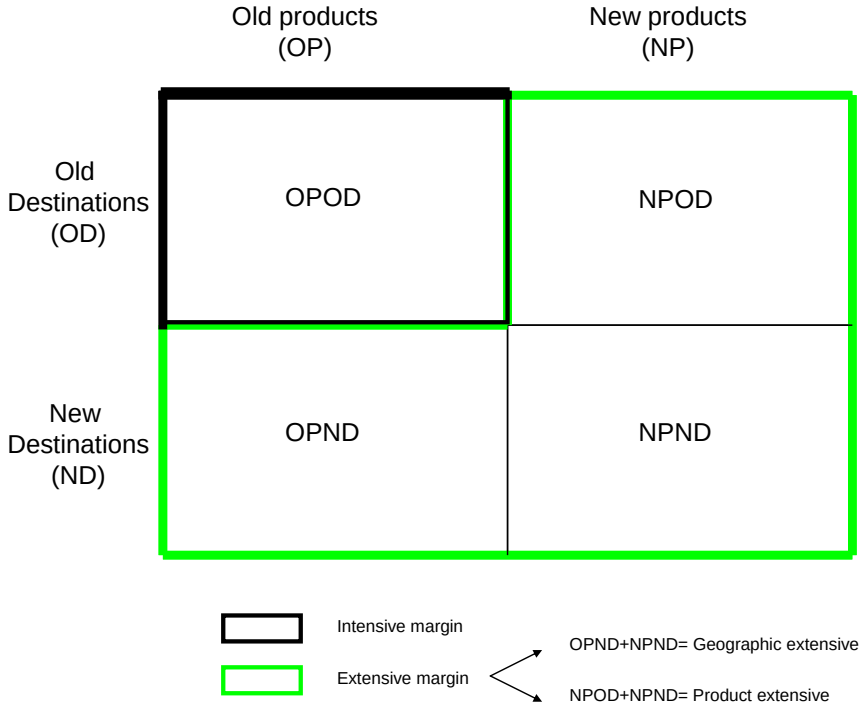
[3] For completeness we use mirror data (imports) from Comtrade.
[4] See Hausmann and Klinger (2006) for a diversification measure considering productivity content.

all products that were exported at least 3 years before 1995, either consecutively or not. Likewise, new products have been defined as those products that have been exported for at least 5 times after 1995.[5]

We consider that this traditional classification is well suited to discuss diversification issues from a product point of view, but it lacks a geographic dimension. This is why we have added a twist to the traditional definition. The twist simply consists in including the geographic dimension in order to allow us to distinguish between product and geographic diversification. This implies that the intensive margin will consist of "Old Products" being exported to "Old Destinations" (OPOD). In the same way, the extensive margin will consist of "Old Products" being exported to "New Destinations" (OPND), "New Products to New Destinations" (NPND), and "New Products to Old Destinations" (NPOD).

To sum up, there are two dimensions to export diversification. Product diversification is the sum of NPND and NPOD, whereas geographical diversification is the sum of NPND and OPND. Figure 1 illustrates our classification. We have classified all products for the set of countries included in our sample for the last 15 years following this classification.

Figure 1: Definitions



Some remarks should be made with respect to this classification. First, it is important to note that there is an overlap in this decomposition of the extensive margin. The

[5] We have followed this criterion to avoid the inconveniences of dealing with noise in the data (many products coming in and out) either before or after the break in 1995. The year 1995 as a breaking point was chosen based on the fact that it would allow us to have some reasonable length of time -10 years- for the entry of new products within our sample period. However, as a robustness check, we also tried with other years as breaking points (1994, 1996, 1997) and the results are very similar.

category New Products to New Destinations (NPND) is included in both, product and geographical diversification. As far as the analysis presented in the paper, this overlap does not affect our findings: it simply implies that the comparison between product and geographical diversification at the extensive margin boils down to compare whether OPND or NPOD is larger than the other.

An additional point to make is on the difference between both margins of trade. As Baldwin (2006) notes, the 6-digit classification is not fine enough to pick up on individual products. There may be some goods that start being traded but that we cannot identify since they fall in categories where some trade is already taking place. The only newly traded goods that can be observed directly are those in categories that switch from zero value to some positive value within our data period. [6] Therefore, for the purpose of our final conclusions, we acknowledge that looking for changes in the number of zeros at this level of disaggregation systematically underestimates the importance of the newly traded goods.

Finally, what is relevant from a product point of view may not be relevant from a geographical one, and vice versa. Traditional classifications of geographic diversification define the extensive margin at the country level. This means that a destination becomes "old" when any of the 6-digit categories is exported to it. The geographic spread can only increase by exporting to completely new countries. However, there is no such a thing as the exports of a “country”. The exporters are firms in different sectors. To overcome this issue and avoid getting a distorted perception of the relevance of geography-driven diversification, we have defined the *geographic extensive* margin at the industry level.[7] We have defined 5-digit "sector-specific destinations", so to speak. An example may help to clarify this point: from a geographic point of view, it is interesting that the furniture industry is exporting cupboards to a wider range of countries regardless of whether those countries were previously recipients of, say, dairy products. But from a geographic point of view it is not very relevant when the furniture industry starts exporting beds to destinations where it was already exporting cupboards. This is precisely the focus of product diversification.

2.3. *Stylized facts on the evolution of export diversification*

We first look at **Figure 2**, which summarizes the evolution of the number of zeros by region.[8] As we can observe, the common feature for the different regions is that the

[6] To be precise with our terminology then, we should be referring to the quasi-intensive and quasi-extensive margins. However, without loss of generality and in order to keep the terminology as simple as possible we drop the term quasi- from the terminology.

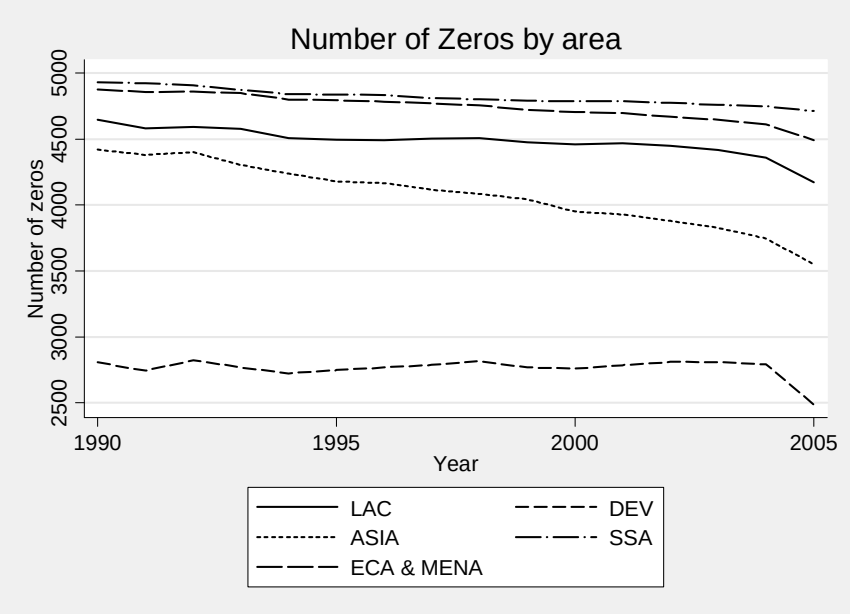
[7] We determine the destination at the HS 5-digit disaggregation level. From the point of view of the trade minister, interested in industrialization in several industries, it matters that the furniture industry is exporting a wider range of furniture varieties. But it is also good the "light bulb" industry starts exporting to a wider range of countries.

[8] We have grouped the countries by regions as follows:
- LAC: Argentina, Brazil, Chile, Peru and Mexico.
- Developed: Australia, Canada, European Union, Japan and the United States.

number of zeros has been declining over the years. This suggests that product diversification is on the rise everywhere. A more careful inspection of the graph suggests that the decline is more substantial for East Asian economies than for the remaining regions. This point is confirmed by the actual data country by country.[9]

The variation is the lowest for developed countries and the group of poorest countries in our sample. This is due to very different reasons. Developed countries display a lower reduction in the number of zeros because there are few 6-digit categories that are not being exported yet. Although by definition all of these categories are tradable, empirical evidence shows that there is some sort of lower bound to the number of zeros. This can be evidenced by intra-European trade (the most integrated trading area in the world) where some categories remain not traded. Further research is needed to understand what the nature of these "rigidities" may be. The poorest countries in our sample, on the other hand, are experiencing a totally different type of constraint. It may be the case that their economic structure is very basic and they concentrate their exports in just a few products (mostly commodities). Another explanation is that they lack the capacity (e.g. knowledge, infrastructure) to innovate (e.g. diversify their production) or simply that the national production does not meet the standards to export.

Figure 2: Evolution of the number of zeros by region (1990 – 2005)



Export diversification seems to be influenced by the proximity and the size to the destination markets.[10] The figures in Appendix 2 show the scatter plots of the count of

- SSA: Ghana, Kenya, Mauritius, South Africa and Uganda.
- ASIA: Bangladesh, China, Indonesia, India, Malaysia and Thailand.
- ECA-MENA: Morocco, Tunisia and Turkey.

Details for individual countries are available from the authors upon request.

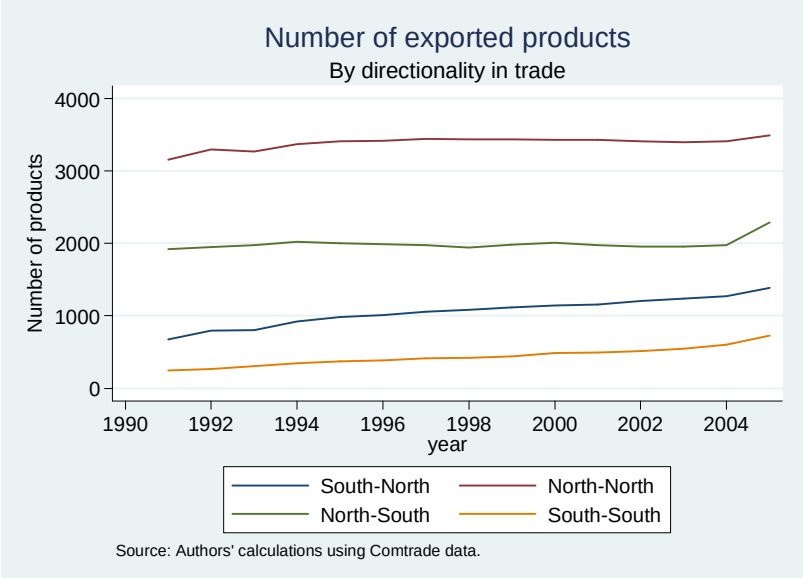
[9] See Table 6 in Appendix 5.

[10] Note that in the previous graph we saw diversification in terms of the evolution of number of zeros, now we switch to observe the evolution in the count of goods exported.

products exported in 2000 by some of the countries in our sample against the GDP of all the destination countries.[11] The graphs show that the number of products exported to markets that are closer and larger is considerably higher than to other destinations. For example, Latin countries such as Argentina and Brazil trade a basket of products significantly wider among themselves and with larger countries like USA, the EEC15, etc. Likewise, East Asian countries trade more goods regionally and with larger markets as expected, however, one point important to note for this group of countries is that in general, their basket of exports to each destination is on average more diversified (more concentration of dots in the upper part of the graphs). Finally, Tunisia and South Africa show similar patterns as those mentioned above, however, in this case, their exports basket is even more concentrated than in Latin countries' case.

In order to check how much the level of development of the origin and destination nations matter in the trade relationship, we look at the evolution of the number of goods traded according to the directionality in trade (whether exports come from or go to the North or the South). Figure 3 shows that the number of products exported over time is growing more for the bilateral relationships involving the South (South-North and South-South). This is quite a natural result if we consider that there are few 6-digit categories that developed countries are not trading among themselves yet. However, as far as the directionality in trade within South, the number of goods traded with the North is considerably higher than that traded with the South.

Figure 3: Evolution of the number of products exported by directionality in trade



Summing up the facts above, diversification is on the rise and it seems to be more dynamic within the developing countries in our sample. The extent to which countries diversify seems to differ depending on the income level, size and proximity of the

[11] We made graphs for years 1990, 1995, 2000 and 2004 and they all give us similar pattern. We only report the scatter graphs for 2000 for simplicity in the presentation of our results, however, graphs for the other years are available upon request.

destination markets.

2.4. *Stylized facts on the margins of trade*

One of the differences between developed and developing countries is that developed countries export a wider range of products to a much wider range of destinations than developing countries do. Moving into this difference with more detail, in our next figure we show the relative distribution of the different margins of trade by region.

Figure 4: Intensive versus Extensive margin by region

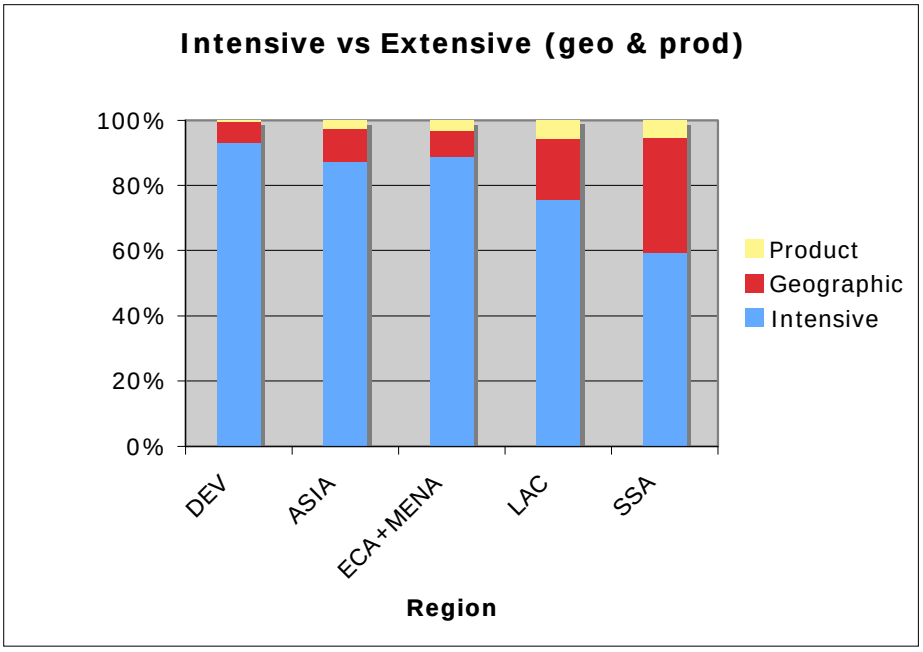


Figure 4 shows that the growth of trade at the intensive margin is much more important than at the extensive margin for all groups of countries. The extensive margin seems to be relatively more important for poorer regions. At the extensive margin geographic diversification is more important than product diversification. It seems that the relative importance of product diversification increases as we move to less developed regions. Countries in Asia and ECA+MENA have similar patterns. ECA+MENA is the region neighboring the EU and the increase may be reflecting outsourcing of production from the EU.[12]

In Table 1 we break down the overall export growth between 1995 and 2005 into the share of the overall growth corresponding to the products that were already being exported (*intensive*) and the growth in the exports of the products that started being exported after 1995 (*extensive*). For expositional simplicity we have aggregated the figures by regions. As it can be observed, the most important share corresponds to the intensive margin. The extensive margin accounts, on average, for 14% of the overall

[12] A similar graph is presented in Appendix 5. It shows a bell-shaped relationship where SSA and LAC show very different pattern than the rest of regions.

export growth whereas the intensive margin on average represents the larger share in the increase. The first figure in Appendix 5 provides the breakdown country and by country. It is worth noting that the amounts vary widely across nations.

Table 1: Growth in exports in new and old products by areas

	Intensive (1)	Extensive (2)		Overall (3)	% Ext (2) / (3)	% Geo. Ext (2.1) / (3)	% Prod. Ext (2.2) / (3)
		Geographic (2.1)	Product (2.2)				
LAC	94.35	23.29	6.97	118.06	0.22	0.20	0.06
DEV	55.60	3.94	0.16	59.59	0.05	0.07	0.00
SSA	28.81	17.00	2.69	45.73	0.16	0.37	0.06
ASIA	133.54	15.59	3.84	150.53	0.14	0.10	0.03
ECA+MENA	104.02	9.36	3.67	114.94	0.12	0.08	0.03

Note: Author's calculations using Comtrade data.

Columns (2.1) and (2.2) decompose the extensive margin into the product and the geographic dimensions as explained in Figure 1.[13] As we can see, geographic diversification is more important than product diversification for all regions.

Finally, before turning to the formal empirical analysis, we explore the sectoral distribution of the extensive margin. With this aim, we have converted the HS-6 categories to the two-digit NACE classification system.[14] After the conversion we are left with roughly thirty categories which we have aggregated for expositional simplicity. We turn now to show what the most important sectors at the extensive margin are. The results are summarized in the following pie-graphs:

Figure 5: New products by sector in developed and developing countries

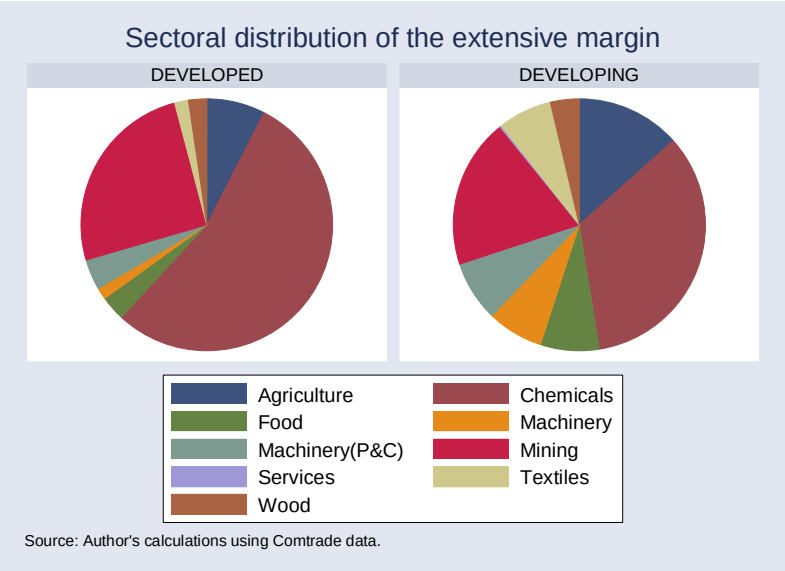


Figure 5 shows the distribution of sectors at the extensive margin in developed versus developing countries. Around half of the new products that the developed countries export are concentrated on chemicals and machinery alone. The figures for mining

[13] Both columns include the category of products defined as NPND, which constitutes the overlap between both definitions. If the category NPND is excluded, the difference between the geographic (OPND) and product (NPOD) dimensions is even larger.
[14] See Appendix 4

suggest that new export products account for a quarter of the value of the extensive margin of trade.

Figure 6: New products by sector and by region in developing countries

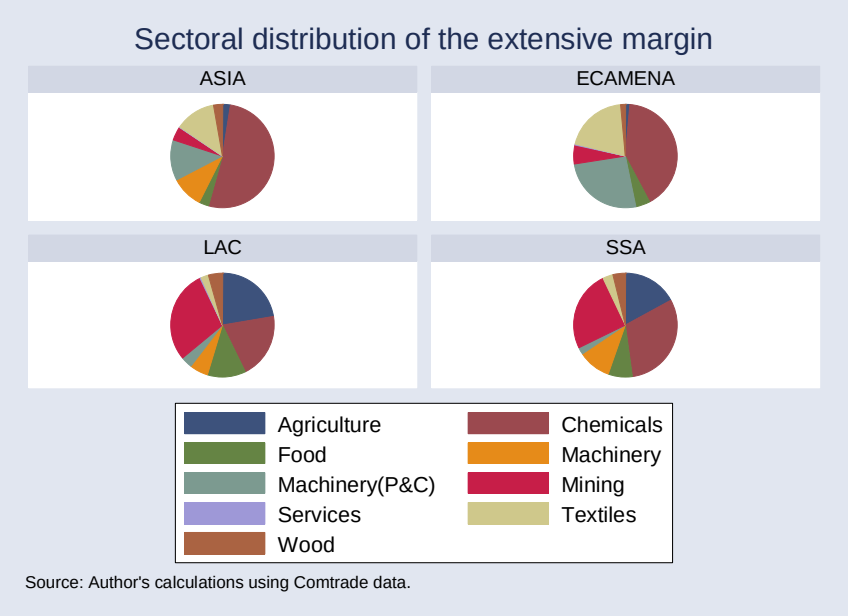


Figure 6 shows the importance of the different sectors at the extensive margin by regions. Several features emerge clearly from these graphs. First, the sectoral distribution of the extensive margin varies among developing countries. In particular, higher income developing countries such as those in ASIA and ECA have a higher share of their extensive margin on machinery than the rest of the countries. Second, agricultural products are more important for the countries in Latin America and Africa than for those in Asia and ECA+ MENA.

To conclude, these charts suggest that there are some clearly defined diversification patterns. That is, the intensive margin is the most important category for all regions; and within the extensive margin, geographical diversification seems to matter more for developing countries than it does for developed ones. Moreover, it seems that the poorer the region, the more important the geographic and product dimensions of the extensive margin. We acknowledge that these are partial results since other factors that affect export diversification are not controlled for in the charts. Undertaking these controls requires more formal theoretical and statistical procedures that will be presented in the following sections.

3. THE THEORY

From the previous section we found that, although the intensive margin explains the bulk of the growth in exports for all countries in the sample, diversification seems to be non-trivial for developing countries and this in turn seems to be determined by–at least- the geographic region where the exporting country is, the size of the destination market and the directionality in trade. Thus, in this section we turn to a presentation of

the theory that we will use to frame our analysis of these elements and that will give foundation to our econometric approach.

Traditional trade theories are ill suited to investigate diversification patterns because they provide no explanation of zero-trade flows. For this reason, we turn to the recent developments in the theory of trade, the so called "new-new trade theory", that has contributed to change this by taking into account the fact that not all firms export. The new paradigm provides a solid theoretical foundation to the investigation of diversification issues by allowing us to deal with zero trade flows in the trade matrix in a natural way.

In particular, we follow the framework based on the seminal paper of Melitz (2003), which allows taking into account firm heterogeneity and market-specific beachhead costs.[15] In this set up, firms try to have large enough sales to make it profitable to cover the sunk costs of entering foreign markets. As a result, the range of firms that export is endogenously determined and related to native firm-level productivity. This accounts for the empirical observation that large firms export while small firms do not.

The equilibrium in nation-*o* (origin) is characterized by a set of cut-off conditions for each market, including its own. The domestic cut-off condition defines the highest marginal cost for which nation-*o* firms still find it worthwhile to produce. The domestic cut-off condition is:

$$F_o^D = \left(\frac{\bar{a}_{oo}}{1-1/\sigma} \right)^{1-\sigma} \frac{\mathbf{B}_0}{\sigma}$$

where F_o^d stands for the cost of entering the domestic market, B_o is the demand shifter in nation-*o*, σ is the elasticity of substitution among varieties, which we assume to be constant and higher than one, and a_{oo}^* is the threshold marginal cost for local sales. We assume negligible trade costs for local sales.

Only firms with sufficiently low marginal costs are able to sell to foreign markets, since only they are able to cover the fixed market-entry costs.[16] Thus, the export threshold determines which domestically produced goods are exported. The export cut-off condition is:

$$F_d^X = \left(\frac{\bar{a}_{od}\tau_{od}}{1-1/\sigma} \right)^{1-\sigma} \frac{\mathbf{B}_d}{\sigma}$$

where F_x is the fixed cost of entering the market in nation-d (destination), a_{od}^* is the pair-specific threshold-marginal-cost (a stands for unit labor input), τ_{od} is the bilateral trade cost, B_d is the demand shifter in nation-d, namely $E_d / p_d^{1-\sigma}$ where E_d is the total expenditure of the destination nation on all varieties, p_d is the usual CES price index, and a_{od} is the threshold marginal cost for sales abroad.

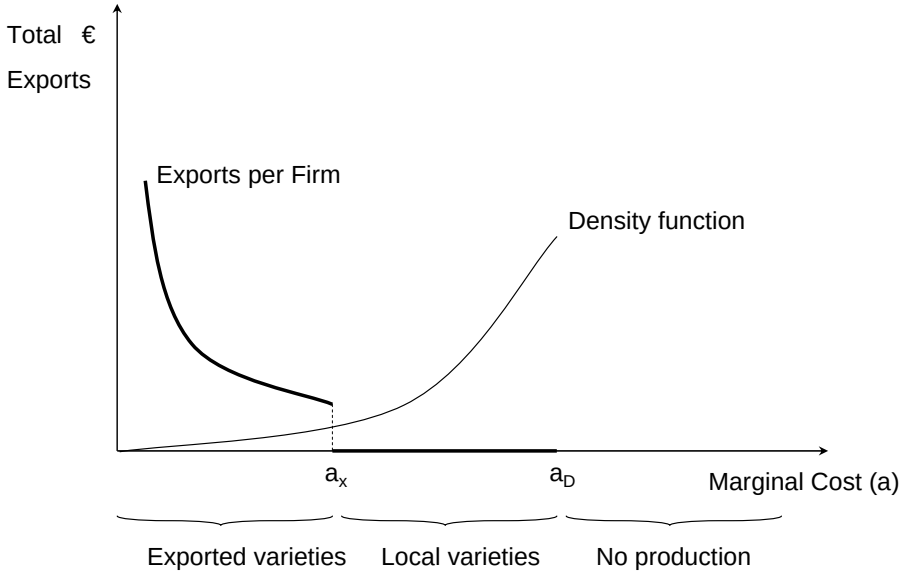
---

[15] This section provides a theoretical background that motivates our empirical work. Interested readers not familiar with the Melitz (2003) model can read the original article, and Baldwin (2005) and Helpman, Melitz and Rubinstein (2007).

[16] Due to the fixed mark-up rule, they fully pass on the per-unit trade costs to foreign customers; the price of their good is higher in foreign markets.

The bilateral exports from nation-o to nation-d are endogenously determined by the domestic and the export cut-off conditions (a_x and a_0 in Figure 7). The economic model consists in determining simultaneously whether to sell and how much. In other words, it is an export decision with a threshold.

Figure 7: Determining the number of goods in a "new-new trade'' model



The second threshold implies that the model's predictions are in line with the common observation that big, efficient firms are more likely to export than small firms. Moreover, the further away is the market, the higher will be the price (due to passed-through trade costs) and so the lower will be the operating profit earned.

The total per-firm value of bilateral exports measured in terms of the numeraire is:

$$V_{od} = \begin{cases} \int_0^{\bar{a}_{od}} n_o \frac{a\tau_{od}}{(1-\frac{1}{\sigma})^{1-\sigma}} \mathbf{B}_d dG[a|a_{oo}^*], if \quad a \le a_{od}^* \\ 0, \quad otherwise. \end{cases} \tag{1}$$

where we have denoted by V_{od} the volume of bilateral exports between the origin and the destination nations. $G\left[a/a_{oo}^*\right]$ is the conditional density function that describes the distribution of marginal costs in nation-o (conditional on the domestic threshold marginal cost a_{oo}^* since firms that do not produce cannot export). Note that, as the threshold marginal cost rises, smaller firms will export their goods, so the range of exported goods will widen.

Re-arranging the variables we have that:

$$V_{od} = \begin{cases} \tau_{od}^{1-\sigma} \mathbf{B}_d (n_o \int_0^{\bar{a}_{od}} a^{1-\sigma} dG[a|a_{oo}^*])(1-\frac{1}{\sigma})^{\sigma-1}, if \quad a \le a_{od}^* \\ 0, \quad otherwise. \end{cases} \tag{2}$$

The expression for the bilateral trade volume, Equation (2), suggests a gravity-like estimation. Noting that B_d equals $E_d / p_d^{1-\sigma}$, we can use the GDP of the importing nation to proxy for E_d, and the GDP of the exporting nation to proxy for n_o; n_o is related to the endowment of the exporting nation[17]. The remaining terms, including bilateral trade costs, $\tau_{od}^{1-\sigma}$, and additional nation-o specific factors affecting n_o can be controlled for using time-invariant pair dummies which in our case, based on the facts observed previously, will pick up on the distance between markets and the directionality in trade (North/South).

This framework is linked to diversification to the extent that it is clear from the model that the range of exported goods is not only affected by the factors we have analyzed before, but also, it is somehow linked to the export threshold. In particular, as the fixed market entry costs have been falling over time for exporting firms, the number of zeros in export vectors should also be falling - fact supported also by the evidence presented in the previous section-. This drop in bilateral trade costs, or the fixed market entry costs, may not only be stimulating bilateral exports, but it can also induce firms to start exporting new categories of goods that were previously not exported or to start selling to new destinations.

One concrete story would be that developing countries have been signing FTAs and as a consequence the associated costs with exporting have been reduced.[18] If the fixed cost of entering the FTA markets falls, then a wider range of firms will find selling to the FTA to be worthwhile. As a consequence, the number of goods exported to the FTA will increase. More goods, in turn, mean a higher trade volume. The fundamental difference between this result and the standard gravity model is that censoring here is an issue. When trading is not profitable, we find a censored data -zero export values- in that category.

Our empirical strategy is to focus on the bilateral trade flows that switch from zero to a positive value for each possible destination, although we also look at the change in existing trade flows. With this aim, in what follows, we undertake the Tobit estimation, which views the non-existent trade flows as censored data.

[17] See Helpman (2004) for details.

[18] For instance, higher transparency in the rules for exporting to FTA members lower information costs for exporters which translates to a fall in the fixed cost of entering the FTA. We are going to assume that the FTA lowers the fixed market entry costs for its members.

4. ECONOMETRIC APPROACH: TOBIT MODEL

We turn now to the bulk of our econometric analysis using the dataset described in Section 2. As we have seen earlier, the fact that some bilateral trade relations switch from zero to positive values (and the factors behind the switch) is what diversification is about. The existence of zeros in the trade matrix is an important piece of information as it accounts for the non-successful experiences of diversification. From an econometric point of view, the way we deal with censored data in our analysis is using the Tobit model.[19] We start the section by presenting our estimating equation, which rests on the theoretical framework in Section 3. Then, we discuss the econometric technique used to estimate the equation (the Tobit model) and we conclude with a discussion of the estimated results.

4.1. *The estimation equation*

Following the general facts previously found and the theoretical framework presented in Section 3, we estimate the following gravity equation:

$$\log(V_{odit}) = \alpha + \beta_1 \log(GDP\_d_t) + \beta_2 \log(Dist_{od}) + \beta_3 Trade\_North + \beta_4 FTA + SD + TD + \varepsilon_t \quad (3)$$

where; V_{odit} is the dollar value of exports from country-o (origin) to country-d (destination), of product i -for each 6-digit category- in year t; GDP_d_t is the GDP of the destination country in year t;[20] $Dist_{od}$ is the distance between country-o and country-d. In principle, we expect the GDP and the distance to show positive and negative sign respectively.[21] As for the additional specific factors in the bilateral relationship, we include the Trade_North dummy that takes the value of 1 when the trade relationship is with a developed country (North).

Also, following the theoretical approach in the previous section and in order to take into account the reduction of trade costs (e.g. the reduction in the market entry costs), we also created the FTA dummy [22], which takes the value of 1 for trade relations between countries parties to the same agreement.[23]

---

[19] The approach that we use in this paper differs from that of Helpman, Melitz, Rubinstein (2007).

[20] We do not include the origin's GDP in the equation because as we will explain later with more detail, due to computational limitations, we have to estimate the equation for each exporter (origin) separately. However, in order to evaluate whether the time-variant GDP of the exporter mattered, we also included the GDP of the exporting country (origin) in alternative estimations, but it did not change the results.

[21] The GDP is expressed in current US dollars and it is extracted from the World Development Indicators. (http://www.worldbank.org). Distance data were obtained from CEPII website (-http://www.cepii.fr/francgraph/bdd/bdd.htm) and it is defined as the distance between the economic centres of one country to another. We have selected Brussels as the economic centre of the EU.

[22] The trade agreements covered with this dummy are those included in the study for the Global Economic Prospects, World Bank (2005), which in turn includes all those agreements notified to the WTO and that can be found at: http://www.wto.org/english/tratop_e/region_e/region_e.htm; and also other agreements not notified.

[23] For robustness purposes we also estimated the model using tariff data -as a proxy of trade costs- for the available 6-digit level categories instead of and in addition to the FTA dummy. Our conclusion is that using tariff data does not undermine the findings. The results do not change drastically.

Finally, since the estimation covers the universe of products exported at the HS 6-digit, we include sector dummies (SD) to control for sectoral differences and time dummies (TD) to deal with the conversion of all the current valued dollars to a common base year. We have included a table in the appendix that summarizes all the variables employed as well as the source of the data.

4.2. *Estimation technique: The Tobit model*

As we have stated before, the core of our econometric analysis of diversification relies on what happens with the trade flows that switch from zero to positive values. Thus, dealing with censored data is at the heart of our work. Several alternatives are available to deal with censored data (the zeros in the trade matrix).

The Poisson Pseudo-Maximum-Likelihood (PPML) estimation and the Tobit model are the most common estimation techniques to deal with the censored data. However, at this point, it is important to recall that, although the focus of our work is diversification (the extensive margin), what happens at the intensive margin (which as we have seen before is very significant for all countries) cannot be ignored for estimation purposes. Therefore, we opted for the Tobit estimation as it gives us flexibility to disentangle what happens at both margins at the same time that it takes into account the censoring structure of the data (which is not the case for the other estimation techniques).[24]

One important point to make about the Tobit model concerns the interpretation of the results. It should be noted that the estimated raw coefficients from the Tobit estimations do not have a particularly interesting economic meaning. These are simply the effect of the independent variables on the "latent" variable that underlies the Tobit model.[25] In order to disentangle both margins and to provide a simple economic interpretation of the parameters we have computed the marginal effects. The marginal effects calculated after the Tobit estimation, show the effects of the independent variables on the overall amount of trade (unconditional expected value), on the intensive margin (conditional marginal effect) and on the probability of exporting more positive values (probability uncensored).[26]

In particular, by analyzing the marginal effects from the Tobit estimation, we pick up the total impact of the different independent variables on both a) the expansion of

---

[24] Alternatively, as a robustness check, we also used a probit model to estimate the marginal effects on the probability to export and we get similar results to those obtained from the Tobit estimation. The tables summarizing these results are available upon request.

[25] See Greene (2003)

[26] This comes from a useful decomposition suggested by Mc Donald and Mofitt (1980) and according to which:

$$\frac{\partial E[y_i \mid x_i]}{\partial x_i} = \Pr ob[y_i > 0]\frac{\partial E[y_i \mid x_i, y_i > 0]}{\partial x_i} + E[y_i \mid x_i, y_i > 0]\frac{\partial \Pr ob[y_i > 0]}{\partial x_i}$$

Therefore, the change in x_i has two effects: it affects the conditional mean in the continuous part of the distribution (the positive values, or in the intensive margin of trade in our case) and it also affects the probability that the observation will fall in that part of the distribution (the switch from zero value to a positive value or the increase in the probability that a country export a more diversified basket of goods). See Greene (2003).

trade in existing categories and b) the change in the probability of exporting more positive values.

However, such computational convenience comes at a price. In a recent paper Silva and Tenreyro (2006) point out that the use of the Tobit model to estimate the gravity equation in the presence of heteroskedasticity may produce inconsistent estimators. To address this problem they suggest the use of the PPML estimation method, as it produces consistent estimators in the presence of heteroskedasticity by under-weighting the observations with bigger variance terms.[27] We have addressed this problem in two ways. First, we estimated the Tobit model using an interval regression with a left censored data (the zeros,) which allows the calculation of a robust variance estimator. We obtained very similar results to those from the standard Tobit method; therefore, to avoid redundancy, we only present the results from the Tobit.

Second, we also estimated the PPML as a robustness check. The results of the raw coefficients from both estimations are presented in Appendix 1. Three observations are relevant to keep in mind: a) the results are similar in signs and significance for the gravity-like variables (distance and GDP), b) although in all variables there are dissimilarities in signs for some countries, the calculations for the averages keep the same signs, and c) the value of the coefficients from the PPML estimation is in general smaller than those obtained from the Tobit estimation; therefore, it should be kept in mind that, although the results for the intensive margin from the Tobit are robust in terms of signs; they could be overestimated.

4.3. *The results*

In this subsection, we present the results of the marginal effects obtained from the Tobit estimations of Equation (3). But before presenting the results, two further points should be noted on the Tobit model. First, since the underlying gravity theory implies the use of logarithms and this may be a problem since we have zero values, we have used a common technique in the empirical trade literature which consists in shifting all trade values by one unit (the variance of the distribution of trade values is not affected but the mean is increased by one unit).

Second, the size of the dataset introduces some computational limitations. For each of our 24 exporters there are 5,019 product categories for each of the 23 potential partners. This adds up to about 115,000 data points per year per exporter. We have 16 years in all, so the dataset is on the order of 1,800,000 data points for each exporting nation. Pooling all 24 exporters together would create a panel of about 43 million data points, a number which is computationally not feasible. To get around this computation problem we are going to use only one exporter's dataset at a time. This means that, we estimate the equation with a Tobit model exporter by exporter. The results of the average marginal effects on the intensive margin, by exporter, are presented in Table 2 (each row represents a separate regression).

[27] To prove their point, they perform a heteroskedasticity test on each of the different results they obtain for the Anderson-van Wincoop gravity model estimated with different techniques (OLS, Tobit, NLS) and find that the PPML is the only model that cannot reject the null hypothesis of homoskedasticity.

Table 2: Tobit estimation – Marginal Effect on the Conditional on being uncensored (intensive margin), by exporter

Dependent variable: log (exports)

	log (GDPd)	log (dist)	North	Parts & Comp	FTA	Constant
LAC						
ARG	0.238***	-0.661***	-0.084***	0.245***	-0.225***	1.466***
	(208.45)	(265.18)	(19.97)	(52.75)	(48.37)	(69.5)
BRA	0.303***	-0.731***	-0.055***	0.451***	0.007	1.321***
	(288.05)	(269.23)	(13.49)	(97.00)	(1.38)	(52.3)
CHL	0.156***	-0.447***	0.130***	0.119***	0.133***	0.300***
	(131.90)	(187.71)	(27.01)	(21.56)	(29.60)	(11.9)
MEX	0.210***	-0.670***	0.015***	0.467***	0.068***	1.883***
	(218.47)	(231.19)	(3.83)	(107.11)	(16.48)	(67.6)
PER	0.142***	-0.343***	0.229***	0.084***	0.543***	-0.153***
	(102.48)	(116.11)	(44.57)	(13.00)	(65.15)	(4.9)
DEVELOPED						
AUS	0.202***	-1.059***	0.290***	0.379***	0.162***	6.087***
	(212.71)	(284.97)	(60.68)	(91.09)	(30.04)	(171.1)
CAN	0.285***	-0.202***	0.044***	0.471***	0.413***	-2.944***
	(297.68)	(132.74)	(11.70)	(116.03)	(120.26)	(128.6)
EEC15	0.523***	-0.319***	0.285***	0.541***	0.817***	-4.128***
	(482.09)	(102.46)	(53.29)	(84.07)	(124.48)	(135.6)
JPN	0.397***	-0.810***	0.128***	0.648***	0.136***	1.159***
	(435.70)	(347.90)	(32.06)	(137.49)	(6.07)	(43.0)
USA	0.637***	-0.341***	-0.061***	0.649***	1.030***	-5.154***
	(626.61)	(134.86)	(13.64)	(116.03)	(138.01)	(187.9)
SSA						
GHA	0.101***	-0.350***	0.284***	0.043***		0.281***
	(53.70)	(62.72)	(34.58)	(5.63)		(5.6)
KEN	0.107***	-0.471***	0.277***	0.071***	0.133***	1.349***
	(85.04)	(172.16)	(52.02)	(14.91)	(22.49)	(81.3)
MUS	0.151***	-0.576***	0.388***	0.080***	0.224***	1.494***
	(87.32)	(76.96)	(55.56)	(13.40)	(26.74)	(26.1)
UGA	0.109***	-0.303***	0.144***	0.106***	-0.022*	-0.398***
	(37.86)	(54.48)	(13.75)	(11.90)	(1.75)	(10.3)
ZAF	0.107***	-1.126***	1.032***	0.275***	0.396***	7.070***
	(108.94)	(208.26)	(213.72)	(60.41)	(67.94)	(170.6)
ASIA						
CHN	0.314***	-0.276***	0.340***	0.067***	0.685***	-3.140***
	(329.14)	(131.21)	(76.84)	(14.24)	(64.50)	(127.4)
IDN	0.203***	-0.391***	0.345***	0.256***	0.256***	-1.084***
	(210.75)	(152.89)	(82.68)	(50.68)	(43.27)	(43.1)
MYS	0.191***	-0.363***	0.375***	0.385***	0.128***	-1.354***
	(205.46)	(155.23)	(89.59)	(83.49)	(22.16)	(55.5)
THA	0.190***	-0.324***	0.520***	0.306***	0.411***	-1.613***
	(214.54)	(168.85)	(125.46)	(66.14)	(83.56)	(73.5)
BGD	0.128***	-0.030***	0.227***	0.037***	0.134***	-3.414***
	(83.90)	(10.54)	(38.61)	(4.76)	(13.76)	(85.2)
IND	0.100***	-0.332***	0.650***	0.359***	-0.064***	-0.145***
	(126.93)	(146.48)	(154.90)	(78.53)	(9.39)	(6.3)
ECA-MENA						
MAR	0.094***	-0.398***	0.279***	0.247***	0.136***	0.843***
	(85.20)	(161.11)	(53.65)	(45.13)	(23.42)	(37.4)
TUN	0.062***	-0.341***	0.377***	0.165***	0.164***	0.479***
	(50.25)	(143.07)	(61.33)	(31.36)	(30.11)	(19.2)
TUR	0.133***	-0.439***	0.334***	0.242***	0.302***	0.464***
	(149.82)	(186.84)	(81.36)	(54.09)	(50.55)	(21.4)

Observations: 1,841,973
Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

As we can see, all coefficients are significant and, although there are differences across countries, the signs obtained for the coefficients of the different independent variables are the same (in most cases). Since we want to analyze differences between

developed and developing countries and in order to present the information in a more organized way, in Table 3 we present a summary of the weighted (by population) average marginal effects of different groupings of countries on the intensive margin.[28]

Table 3: Tobit estimation – Average Marginal Effect on the intensive margin (by groupings)

Dependent variable: log (exports)

	log (GDPd)	log (dist)	North	Parts & Comp	FTA
LAC	0.252	-0.665	-0.009	0.392	0.046
Developed	0.527	-0.412	0.134	0.587	0.759
SSA	0.107	-0.646	0.516	0.146	0.171
ASIA	0.210	-0.296	0.458	0.200	0.326
ECA - MENA	0.116	-0.419	0.323	0.236	0.245

Authors calculations using COMTRADE and WDI data. The estimation also includes sector and time dummies.

Groupings: LAC: Argentina, Brazil, Chile, Peru and Mexico. Developed: Australia, Canada, European Union, Japan and the United States. SSA: Ghana, Kenya, Mauritius, South Africa and Uganda. ASIA: Bangladesh, China, Indonesia, India, Malaysia and Thailand. ECA-MENA: Morocco, Tunisia and Turkey.

The results obtained for the gravity variables reflect that the larger the size of the market at destination and the closer the markets (lower trade costs), the larger the increase in the volume of exports. This finding supports the evidence presented in Appendix 2. Distance seems to be, on average, more important for the sample of developing countries; and within our group of developing countries, we observe that the coefficients are larger for Latin American and African countries. On the other hand, the size of the destination market is more important than distance for Developed, Latin American and Asian countries.

As for the North dummy, as we would expect, trading with the North is more important for developing countries, in particular for countries in Africa and Asia. This is in line with the stylized facts presented in section 2.

Signing an FTA and therefore reducing trade costs has a positive effect on the intensive margin as we would expect and significant in the majority of the cases. As for the groupings, we can observe that developed countries seem to be the ones benefiting more from this, followed by the Asian countries (that trade more with the North) and the countries surrounding Europe (ECA-MENA).

Given that trade in parts and components is becoming more important in recent years for some developing countries, we report the coefficient obtained for the dummy for this group of products. As we can see, we obtain positive coefficients confirming the premium that trade in parts and components has for developing countries. Although the average coefficient is higher for developed countries, the countries in Latin America, Asia and ECA-MENA regions (groups highly integrated to large markets like U.S. and Europe) show also significant premiums.

As for the marginal effects on the change of the probability of observing more positive values (exporting more goods), the results are presented in Table 4.

[28] The source of the data on population is the World Development Indicators.

Table 4: Tobit estimation – Marginal Effect on the change in the probability to export more goods (extensive margin), by exporter

Dependent variable: log (exports)

	log (GDPd)	log (dist)	North	Parts & Comp	FTA	Constant
LAC						
ARG	0.024***	-0.067***	-0.008***	0.029***	-0.019***	0.148***
	(208.45)	(265.18)	(19.97)	(52.75)	(48.37)	(69.51)
BRA	0.055***	-0.134***	-0.010***	0.090***	0.001	0.242***
	(288.05)	(269.23)	(13.49)	(97.00)	(1.38)	(52.30)
CHL	0.011***	-0.033***	0.010***	0.009***	0.011***	0.022***
	(131.90)	(187.71)	(27.01)	(21.56)	(29.60)	(11.85)
MEX	0.030***	-0.095***	0.002***	0.079***	0.010***	0.266***
	(218.47)	(231.19)	(3.83)	(107.11)	(16.48)	(67.56)
PER	0.005***	-0.013***	0.010***	0.003***	0.037***	-0.006***
	(102.48)	(116.11)	(44.57)	(13.00)	(65.15)	(4.91)
DEVELOPED						
AUS	0.043***	-0.225***	0.065***	0.087***	0.036***	1.293***
	(212.71)	(284.97)	(60.68)	(91.09)	(30.04)	(171.06)
CAN	0.069***	-0.049***	0.011***	0.120***	0.104***	-0.710***
	(297.68)	(132.74)	(11.70)	(116.03)	(120.26)	(128.64)
EEC15	0.084***	-0.051***	0.044***	0.077***	0.109***	-0.666***
	(482.09)	(102.46)	(53.29)	(84.07)	(124.48)	(135.61)
JPN	0.112***	-0.229***	0.036***	0.171***	0.038***	0.328***
	(435.70)	(347.90)	(32.06)	(137.49)	(6.07)	(42.98)
USA	0.152***	-0.081***	-0.015***	0.135***	0.196***	-1.228***
	-626.61	-134.86	-13.64	-116.03	-138.01	-187.85
SSA						
GHA	0.001***	-0.003***	0.004***	0.000***		0.003***
	(53.70)	(62.72)	(34.58)	(5.63)		(5.61)
KEN	0.004***	-0.017***	0.014***	0.003***	0.006***	0.050***
	(85.04)	(172.16)	(52.02)	(14.91)	(22.49)	(81.29)
MUS	0.002***	-0.008***	0.009***	0.001***	0.005***	0.022***
	(87.32)	(76.96)	(55.56)	(13.40)	(26.74)	(26.05)
UGA	0.000***	-0.001***	0.001***	0.001***	-0.000*	-0.002***
	(37.86)	(54.48)	(13.75)	(11.90)	(1.75)	(10.31)
ZAF	0.018***	-0.185***	0.202***	0.049***	0.074***	1.162***
	(108.94)	(208.26)	(213.72)	(60.41)	(67.94)	(170.63)
ASIA						
CHN	0.084***	-0.074***	0.089***	0.018***	0.169***	-0.843***
	(329.14)	(131.21)	(76.84)	(14.24)	(64.50)	(127.36)
IDN	0.033***	-0.064***	0.060***	0.045***	0.045***	-0.178***
	(210.75)	(152.89)	(82.68)	(50.68)	(43.27)	(43.14)
MYS	0.032***	-0.061***	0.067***	0.071***	0.022***	-0.228***
	(205.46)	(155.23)	(89.59)	(83.49)	(22.16)	(55.48)
THA	0.038***	-0.064***	0.109***	0.064***	0.087***	-0.319***
	-214.54	-168.85	-125.46	-66.14	-83.56	-73.51
BGD	0.002***	-0.001***	0.005***	0.001***	0.003***	-0.061***
	(83.90)	(10.54)	(38.61)	(4.76)	(13.76)	(85.21)
IND	0.024***	-0.081***	0.162***	0.090***	-0.015***	-0.035***
	(126.93)	(146.48)	(154.90)	(78.53)	(9.39)	(6.25)
ECA-MENA						
MAR	0.004***	-0.015***	0.014***	0.013***	0.006***	0.032***
	(85.20)	(161.11)	(53.65)	(45.13)	(23.42)	(37.38)
TUN	0.001***	-0.007***	0.013***	0.005***	0.005***	0.010***
	-50.25	-143.07	-61.33	-31.36	-30.11	-19.15
TUR	0.017***	-0.056***	0.048***	0.035***	0.045***	0.059***
	(149.82)	(186.84)	(81.36)	(54.09)	(50.55)	(21.44)

Observations: 1,841,973
Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Again, we get all coefficients significant as with the intensive margin and, for simplicity, we present a summary table of the coefficients by groups of countries in Table 5. In particular, this is the set of results more relevant for the purpose of the focus of our study as it deals specifically with the extensive margin of trade (the probability that zero values switch to be positive values).

Table 5: Tobit estimation – Average Marginal Effect on the change in the probability to export more goods (by groupings)

	Dependent variable: log (exports)				
	log (GDPd)	log (dist)	North	Parts & Comp	FTA
LAC	0.039	-0.103	-0.004	0.071	0.005
Developed	0.110	-0.092	0.022	0.113	0.126
SSA	0.008	-0.070	0.076	0.018	0.028
ASIA	0.051	-0.072	0.111	0.048	0.077
ECA - MENA	0.012	-0.041	0.036	0.026	0.031

Authors calculations using COMTRADE and WDI data. The estimation also includes sector and time dummies.

Groupings: LAC: Argentina, Brazil, Chile, Peru and Mexico. Developed: Australia, Canada, European Union, Japan and the United States. SSA: Ghana, Kenya, Mauritius, South Africa and Uganda. ASIA: Bangladesh, China, Indonesia, India, Malaysia and Thailand. ECA-MENA: Morocco, Tunisia and Turkey.

As we can see, we confirm the evidence found before with respect to the importance of the gravity pattern in the determination of a change in the probability to export more goods. Distance and size of the destination market determine significantly the change in the probability to diversify. Again, we observe a higher impact of destination market size for developed countries and of distance for developing countries.

Trading with a partner in the North also helps to the process of diversification. The impact is slightly larger for developing countries, especially for Asian, African and ECA-MENA countries.

Signing FTAs and therefore reducing trade costs helps to boost diversification. Developed countries benefit more from this. Of the developing countries, those who seem to benefit the most from this are the Asian countries (trade more with the North) and the countries surrounding Europe (ECA-MENA). This finding also emphasizes the importance of the gravity variables in our estimation, as it would show that reducing trade costs impacts more on trade when the agreement is either with a larger or a closer economy. Finally, trade in parts and components also helps the diversification process, particularly in the case of Asian and Latin American countries.

Even though these last results do not provide estimation on the exact impact on the extensive margin of trade, they provide an indication of how the likelihood to diversify more is affected and confirm the main finding about the impact of our gravity and trade costs variables in creating trade in new varieties.

5. CONCLUDING REMARKS AND POLICY RECOMMENDATIONS

This paper contributes to the export diversification literature by presenting and discussing new stylized facts on the differences in diversification patterns between developed and developing countries. We find that export growth is mostly explained by the growth at the intensive margin. This is the situation for all countries in our sample, but very particularly for developed countries where the extensive margin of trade is minimal at this level of disaggregation. Despite that, we also find that diversification is on the rise among developing countries. In fact, the extensive margin of trade is not trivial for some developing countries, and in particular, when we analyze two different dimensions of export diversification (product and geographical) we find that geographical diversification shows more dynamism than product diversification.

Another contribution of the paper relies on the investigation of how export diversification reacts to policy. For this purpose, we rely on a theoretical framework that explains export diversification. In fact, using the Melitz trade model, we have estimated the gravity equation with a Tobit technique. This has allowed us to go deeper into the determinants of the margins of trade. We find strong empirical evidence suggesting that the increase in the number of traded varieties follows a gravity-like pattern in which geographical factors (proxied by distance) and the size of the market of the destination nation significantly determine the change in the probability that a country exports a more diversified basket of goods. We also find that trading with a partner in the North increases the probability to export more goods especially for developing countries. Likewise, we also find that signing FTAs and therefore reducing trade costs helps to boost diversification. This last finding would indicate that reducing trade costs, on average, increases the chances of exporting a wider variety of goods. We believe that these results are not conclusive but increase our ability to associate the reduction of the zeros (or the increase in diversification) with the effects of liberalization.

From a policy perspective, these results point quite clearly at some policy areas where policy-makers could engage in order to promote trade and diversification. The finding that most of the export growth takes place at the intensive margin could be used by governments with scarce resources to foster export promotion activities rather than to focus on innovation on the basis of a higher expected payoff. For instance, moving up the quality chain in the existing exports seems to be more important. Policy-makers should especially take into account that at the extensive margin, geographical diversification is more important than product diversification. In other words, focusing on product innovation may not necessarily always be the best course of action.

Another area where policy-makers can help has to do with the reduction of trade costs. We have seen that the variables capturing trade costs (distance and FTA) have a significant impact in the determination of the intensive margin and the changes in the extensive margin. Although one cannot do anything about the geographical distance between countries, there are policy actions that can be taken to reduce the importance of these costs. For example, such actions could include helping companies to gain market access, promoting effective regionalism, and emphasizing efforts in trade facilitation and infrastructure cost reduction.

Another policy recommendation has to do with the relative importance of trade in parts and components, in particular in Asia and Latin America, and what this implies for the process of diversification. In fact, one of the consequences of the fragmentation of production is that nowadays, manufactures are no longer produced in a single location, on the contrary, bits and pieces travel the world before being assembled and this is becoming important for developing countries. This process is already showing up and has been analyzed in previous studies by Ando & Kimura (2005). Our results confirm evidence along this line: as we saw from the analysis of the stylized facts and from the estimations, trade in these goods explains an important share of the change of the extensive margin in machinery (51.65% for developing countries), which in turn is the second sector with the largest increase in the extensive margin.

From a policy perspective, this could mean that the development of a particular industry (in the traditional way: producing final goods) is not a necessary condition for developing countries to succeed in their pursuit of a boost in diversification. Exploiting their labor capacities in producing parts could be equally important to achieve that goal. However, more research is required in this arena and in particular, it would be important to address more detailed issues like the implications of this type of trade in terms of employment, spillovers for the economy and growth.

Finally, further research could deal with the econometrics of the zeros. The presence of zeros in gravity estimations is quite significant and it increases with the level of disaggregation of the dataset. This imposes serious computational constraints on the estimating techniques that can be employed. Further empirical work is also needed to completely sort out the true nature of the zeros; are they missing observations? Which ones are the result of rounding up small amounts and which ones indicate the absence of trade? It would be very useful to know whether there is any kind of systematic pattern in the zeros.

REFERENCES

Ali, Ridwan, Jeffrey Alwang and Paul B. Siegel 1991. "Is Export Diversification the Best Way to Achieve Export Growth and Stability?." The World Bank. Policy, Research and External Affairs, Working Paper, WPS 729.

Al-Marhubi, Fahim, 2000. "Export Diversification and Growth: An Empirical Investigation," Applied Economics Letters, Taylor and Francis Journals, vol. 7(9), pages 559-62, September.

Amurgo-Pacheco, Alberto 2006. "Preferential Trade Liberalization and the Range of Exported Products: The Case of Euro-Mediterranean FTA". Ph.D. Dissertation, Graduate Institute of International Studies.

Ando & Kimura. 2005. "Two dimensional fragmentation in East Asia: Conceptual Framework and Empirics". International Review of Economics and Finance. 14. Issue 3: 317-348

Baldwin, Richard and Frederic Robert-Nicoud. 2005. "Trade and Growth with Heterogenous Firms". CEPR Discussion Paper No. 4965

Baldwin, Richard 2005. "Heterogeneous Firms and Trade: Testable and Untestable Properties of the Melitz Model", Working Paper No. 11471, NBER.

Baldwin, Richard and Virginia Di Nino 2006. "Euros and zeros: The common currency effect on trade in new goods". HEI Working Paper 21-2006.

Baldwin, Richard, 2006. "The euro's trade effect," Working Paper Series 594, European Central Bank.

Bernard, Andrew and J. Bradford Jensen. 1999. "Exceptional exporter performance: cause effect or both?" *Journal of International Economics*, 47(1):1-25.

Bonaglia, Federico and Kiichiro Fukasaku 2003. "Export Diversification in Low-Income Countries: An international challenge after Doha". OECD Development Centre. Technical Papers 209.

Easterly and Kraay, 2000, "Small States, Small Problems? Income, Growth, and Volatility in Small States." *World Development*, 28(11):2013-27.

Erdem, Erkan, and James Tybout. 2003. "Trade Policy and Industrial Sector Responses: Using Evolutionary Models to Interpret the Evidence." NBER Working Paper 9947, National Bureau of Economic Research, Cambridge, Mass.

Evenett, Simon and Anthony Venables 2002. "Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows," University of Bern. working paper, mimeo.

Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices" American Economic Review, 84(1):157—177.

Greene, W. H. 2003. "Econometric Analysis. Prentice Hall, fifth edition edn.

Grossman, Gene M. & Elhanan Helpman.

- 1991a. "*Trade, knowledge spillovers, and growth*, European Economic Review, Elsevier, vol. 35(2-3), pages 517-526.

- 1991b. “Innovation and Growth in the World Economy”. Cambridge, MA. MIT Press.

Hausmann, Ricardo, and Dani Rodrik. 2003. “Economic development as self-discovery.” *Journal of Development Economics*, 72:603-633.

Hausmann, Ricardo and Bailey Klinger, 2006. “Structural Transformation and Patterns of Comparative Advantage in the Product Space”CID Working Paper No. 128, August 2006

Helpman, Elhanan, Marc Melitz and Yona Rubinstein. 2007. “Estimating Trade Flows: Trading Partners and Trading Volumes,” Harvard University working paper, mimeo.

Herzer, Dierk and Felicitas Nowak-Lehmann D. 2004. “Export Diversification, Externalities and Growth”. Ibero-America Institute for Economic Research. Discussion Paper 99.

Imbs, Jean, and Romain Wacziarg. 2003. “Stages of Diversification.” *American Economic Review*, March, 93(1):63-86.

Jansen, Marion. 2004. “Income Volatility in Small and Developing Economies: Export Concentration Matters.” Discussion Paper 3. World Trade Organization, Geneva.

Klinger, Bailey, and Daniel Lederman. 2004. “Discovery and Development: an Empirical Exploration of ‘New’ Products,” World Bank Policy Research Working Paper 3450., World Bank, Washington, D.C.

Klinger, Bailey and Lederman, Daniel, 2005: *Diversification, Innovation, and Imitation off the Global Technological Frontier*, The World Bank, mimeo

Lederman Daniel and William Maloney. 2002. “Trade Structure and Growth”. The World Bank.

Newfarmer Richard and Paul Brenton (2007). Watching more than the Discovery Channel: Export Cycles and Diversification in Development.

Melitz M. 2003. “The impact of trade in intra-industry reallocations and aggregate industry productivity”. Econometrica 71(6), 1695-1725.

Rodrik, Dani. 2004. “Industrial Policy for the Twenty-First Century.” Centre for Economic Policy Research Discussion Paper 4767. CEPR, London.

Silva, Santon and Silvana Tenreyro (2006). The Log of Gravity. The Review of Economics and Statistics, MIT Press, vol. 88(4), pages 641-658, 09.

Tybout, James R. 2000. "Manufacturing firms in Developing countries: How well do they do and why? *Journal of Economic Literature*. 38(1):11-44.

Tybout, James R. 2001. "Plant- and Firm-Level Evidence on "New" Trade Theories" NBER Working Papers 8418, National Bureau of Economic Research, Inc.

UNESCAP. 2004. “Export Diversification and Economic Growth: The Experience of Selected Least Developed Countries”. Development Papers Series, No. 24.

World Bank, 2005. Global Economic Prospects 2005: Trade, Regionalism and Development.

Appendix 1a

Tobit estimation – Raw Coefficients

Dependent variable: log (exports)

	log (GDPd)	log (dist)	Trade_North	Parts & Comp	FTA	Constant
ARG	1.871***	-5.191***	-0.666***	1.799***	-1.894***	11.510***
	(208.45)	(265.18)	(19.97)	(52.75)	(48.37)	(69.51)
AUS	1.109***	-5.811***	1.509***	1.906***	0.852***	33.409***
	(212.71)	(284.97)	(60.68)	(91.09)	(30.04)	(171.06)
BGD	1.561***	-0.368***	2.645***	0.449***	1.562***	-41.729***
	(83.90)	(10.54)	(38.61)	(4.76)	(13.76)	(85.21)
BRA	1.738***	-4.198***	-0.320***	2.346***	0.043	7.587***
	(288.05)	(269.23)	(13.49)	(97.00)	(1.38)	(52.30)
CAN	1.455***	-1.032***	0.223***	2.158***	1.932***	-15.026***
	(297.68)	(132.74)	(11.70)	(116.03)	(120.26)	(128.64)
CHL	1.332***	-3.806***	1.079***	0.984***	1.099***	2.554***
	(131.90)	(187.71)	(27.01)	(21.56)	(29.60)	(11.85)
CHN	1.063***	-0.934***	1.105***	0.226***	2.040***	-10.630***
	(329.14)	(131.21)	(76.84)	(14.24)	(64.50)	(127.36)
EEC15	0.993***	-0.605***	0.528***	0.971***	1.433***	-7.835***
	(482.09)	(102.46)	(53.29)	(84.07)	(124.48)	(135.61)
GHA	1.396***	-4.822***	3.670***	0.583***		3.870***
	(53.70)	(62.72)	(34.58)	(5.63)		(5.61)
IDN	1.216***	-2.347***	1.962***	1.448***	1.449***	-6.497***
	(210.75)	(152.89)	(82.68)	(50.68)	(43.27)	(43.14)
IND	0.468***	-1.545***	2.759***	1.553***	-0.304***	-0.676***
	(126.93)	(146.48)	(154.90)	(78.53)	(9.39)	(6.25)
JPN	1.381***	-2.814***	0.437***	2.010***	0.459***	4.027***
	(435.70)	(347.90)	(32.06)	(137.49)	(6.07)	(42.98)
KEN	1.179***	-5.203***	2.845***	0.767***	1.389***	14.883***
	(85.04)	(172.16)	(52.02)	(14.91)	(22.49)	(81.29)
MAR	1.003***	-4.257***	2.797***	2.440***	1.377***	9.021***
	(85.20)	(161.11)	(53.65)	(45.13)	(23.42)	(37.38)
MEX	1.437***	-4.590***	0.105***	2.849***	0.460***	12.900***
	(218.47)	(231.19)	(3.83)	(107.11)	(16.48)	(67.56)
MUS	1.955***	-7.449***	4.582***	1.008***	2.671***	19.331***
	(87.32)	(76.96)	(55.56)	(13.40)	(26.74)	(26.05)
MYS	1.131***	-2.146***	2.094***	2.099***	0.738***	-8.014***
	(205.46)	(155.23)	(89.59)	(83.49)	(22.16)	(55.48)
PER	1.482***	-3.586***	2.285***	0.853***	4.784***	-1.598***
	(102.48)	(116.11)	(44.57)	(13.00)	(65.15)	(4.91)
THA	1.031***	-1.760***	2.612***	1.558***	2.046***	-8.764***
	(214.54)	(168.85)	(125.46)	(66.14)	(83.56)	(73.51)
TUN	0.755***	-4.168***	4.184***	1.900***	1.868***	5.843***
	(50.25)	(143.07)	(61.33)	(31.36)	(30.11)	(19.15)
TUR	0.972***	-3.206***	2.285***	1.654***	2.012***	3.383***
	(149.82)	(186.84)	(81.36)	(54.09)	(50.55)	(21.44)
UGA	1.708***	-4.765***	2.190***	1.609***	-0.350*	-6.259***
	(37.86)	(54.48)	(13.75)	(11.90)	(1.75)	(10.31)
USA	1.541***	-0.825***	-0.149***	1.437***	2.172***	-12.469***
	(626.61)	(134.86)	(13.64)	(116.03)	(138.01)	(187.85)
ZAF	0.671***	-7.051***	5.478***	1.612***	2.235***	44.274***
	(108.94)	(208.26)	(213.72)	(60.41)	(67.94)	(170.63)
Averages	**1.269**	**-3.437**	**1.927**	**1.509**	**1.308**	**2.216**

Due to computational limitations, each row represents a separate estimation.
Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

This estimation includes time and year effects. Source: Authors calculations using COMTRADE data.

Appendix 1b

Poisson estimation – Raw coefficients

Dependent variable: log (exports)

	log (GDPd)	log (dist)	Trade_North	Parts & Comp	FTA	Constant
ARG	0.799***	-0.849***	-0.69	0.292**	0.575	3.237
	(6.78)	(3.47)	(1.41)	(2.38)	(1.26)	(1.17)
AUS	0.829***	-2.124***	0.034	0.600***	-0.044	16.749***
	(7.79)	(7.75)	(0.10)	(7.43)	(0.15)	(6.66)
BGD	1.249***	0.184	0.846	-0.07	1.194	-23.200***
	(7.22)	(0.27)	(1.21)	(0.10)	(1.05)	(3.50)
BRA	0.992***	-0.671***	-0.604**	0.500***	0.986***	-0.286
	(13.18)	(3.05)	(2.23)	(4.94)	(2.58)	(0.13)
CAN	0.856***	-0.816***	0.058	0.228**	0.248*	2.994
	(6.69)	(7.45)	(0.14)	(2.33)	(1.72)	(1.35)
CHL	0.977***	-0.152	-0.685**	0.027	0.194	-5.450***
	(8.67)	(0.93)	(2.26)	(0.15)	(1.25)	(3.01)
CHN	0.900***	-0.101*	0.606***	-0.21	1.045***	-6.069***
	(21.67)	(1.95)	(2.77)	(1.13)	(2.96)	(5.91)
EEC15	0.724***	-0.722***	0.03	0.439***	0.168	4.927***
	(11.46)	(3.70)	(0.12)	(7.18)	(0.92)	(3.86)
GHA	0.903***	-2.382***	0.458	-1.620***		11.865**
	(5.16)	(4.36)	(0.64)	(23.01)		(2.19)
IDN	0.991***	-1.273***	0.430***	0.099	0.228*	5.142***
	(35.36)	(17.25)	(3.56)	(0.37)	(1.85)	(5.31)
IND	0.902***	-0.562*	-0.047	-0.520*	1.033*	-1.69
	(8.59)	(1.83)	(0.10)	(1.86)	(1.67)	(0.48)
JPN	0.757***	-0.891***	0.542*	0.628***	0.414	2.768
	(6.36)	(5.83)	(1.88)	(4.38)	(1.60)	(1.15)
KEN	1.133***	-3.013***	1.012	-0.021	0.427**	11.328***
	(5.67)	(8.95)	(1.59)	(0.05)	(2.57)	(13.62)
MAR	0.946***	-1.705***	-1.146	2.206***	0.735***	6.202**
	(7.90)	(4.84)	(1.47)	(24.39)	(3.56)	(2.11)
MEX	1.072***	-2.429***	-0.614***	0.407***	0.277***	13.766***
	(39.26)	(37.16)	(3.07)	(13.47)	(2.94)	(18.26)
MUS	2.490**	-4.664***	0.147	-0.592***	2.08	6.524
	(2.13)	(4.39)	(0.06)	(7.89)	(1.55)	(0.79)
MYS	0.973***	-0.348**	0.054	0.896***	1.013***	-3.961**
	(9.53)	(2.18)	(0.15)	(4.56)	(2.63)	(1.99)
PER	1.243***	-0.249	-1.231**	-0.853**	2.800***	-8.982**
	(6.38)	(1.34)	(1.99)	(2.13)	(5.52)	(2.42)
THA	0.894***	-0.565***	0.604**	0.354	1.221***	-3.494
	(13.17)	(3.29)	(2.06)	(1.51)	(3.56)	(1.61)
TUN	1.053***	-2.075***	-1.124	1.773***	0.944***	6.402**
	(7.15)	(8.87)	(1.37)	(24.44)	(5.32)	(2.28)
TUR	0.838***	-1.638***	0.483	0.018	0.266	5.665***
	(7.55)	(8.27)	(1.07)	(0.37)	(1.51)	(2.70)
UGA	1.395***	-3.009***	1.264	0.19	-0.974	4.353
	(3.41)	(5.46)	(0.91)	(0.84)	(1.24)	(1.29)
USA	0.888***	-0.486***	-0.34	0.962***	0.996**	0.391
	(8.47)	(3.35)	(0.93)	(14.80)	(2.33)	(0.25)
ZAF	0.780***	-1.638**	0.648	-0.107	0.628**	10.599*
	(6.72)	(2.33)	(1.26)	(0.71)	(2.29)	(1.76)
Averages	**1.024**	**-1.341**	**0.031**	**0.234**	**0.715**	**2.491**

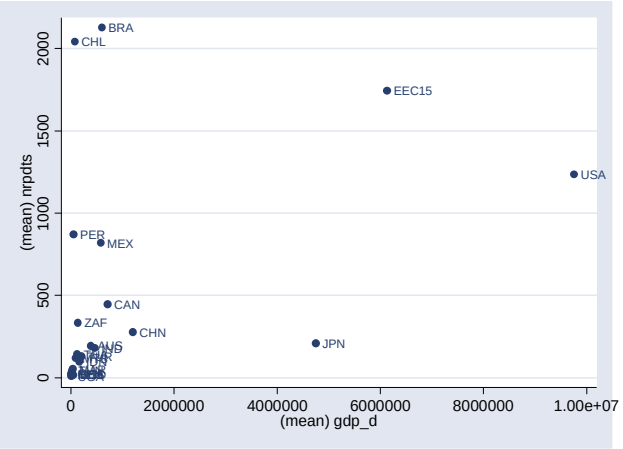
Due to computational limitations, each row represents a separate estimation.
Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

This estimation includes time and year effects. Source: Authors calculations using COMTRADE data.

Appendix 2

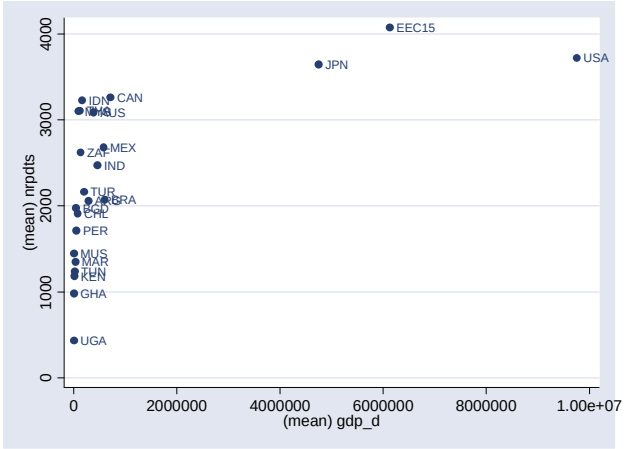
Scatter graph of the count of the number of products exported (Y axis) from origin against the GDP of the destinations (X axis), 2000

Origin: Argentina



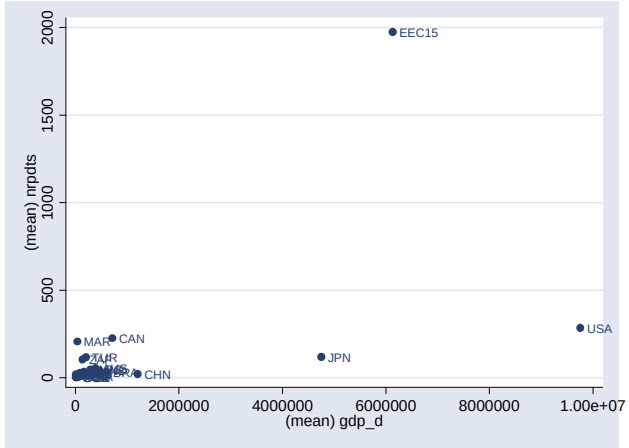
Origin: Brazil

Origin: China

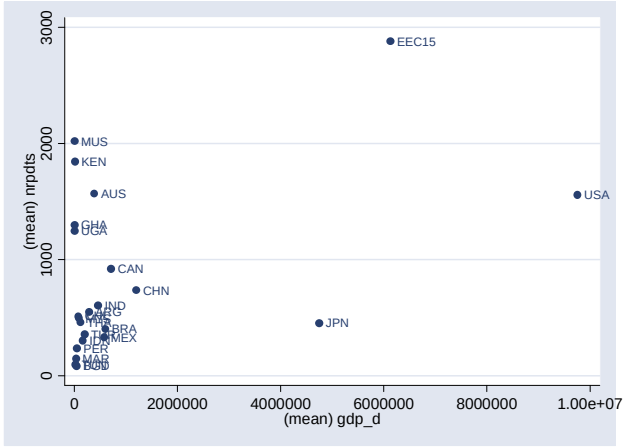


Origin: Malaysia

Origin: Tunisia



Origin: South Africa



Source: Authors calculations using COMTRADE and WDI data.

Appendix 3

Description of the main variables used and their sources

VARIABLE	DESCRIPTION	SOURCE
V_{odit}	Dollar value of exports from country-o (origin) to country-d (destination), of product i -for each 6-digit category- in year t; Using mirror data. Trade flows at 6-digit of the Harmonised System	COMTRADE - WITS
GDP_d_t	Current GDP of the destination country in year t in US dollars	World Development Indicators – World Bank
$Dist_{od}$	Distances between the major cities in each of the countries within the sample	Centre d'Etudes Prospectives et d'Informations Internationales
FTA	Dummy variable which takes the value of 1 for the regional trade agreements (Free Trade Agreements and Customs Unions) between countries parties to the same agreement.	World Trade Organisation and Global Economic Prospects 2005
Other data used for other calculations		
Population	Total population year 2005	World Development Indicators – World Bank

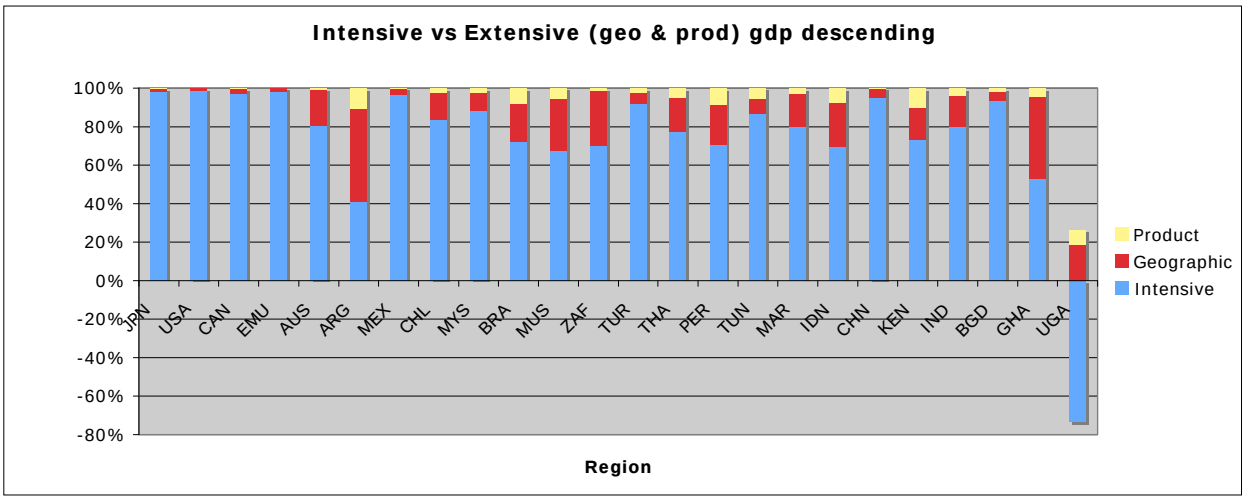
Appendix 4: Sectors according to NACE Classification

Details of the aggregation process:

Nace	Description
	AGRICULTURE
01	Agriculture, hunting and related service activities;
02	Forestry, logging and related service activities;
05	Fishing, operation of fish hatcheries and fish farms; service activities;
	MINING
10	Mining of coal and lignite; extraction of peat;
11	Extraction of crude petroleum and natural gas; service activities;
12	Mining of uranium and thorium ores;
13	Mining of metal ores;
14	Other mining and quarrying;
	FOOD
15	Manufacture of food products and beverages;
16	Manufacture of tobacco products;
	TEXTILES
17	Manufacture of textiles;
18	Manufacture of wearing apparel; dressing and dyeing of fur;
19	Tanning and dressing of leather; manufacture of luggage, handbags;
	WOOD
20	Manufacture of wood and of products of wood and cork, except furniture;
21	Manufacture of pulp, paper and paper products;
22	Publishing, printing and reproduction of recorded media;
	CHEMICALS
23	Manufacture of coke, refined petroleum products and nuclear fuel;
24	Manufacture of chemicals and chemical products;
25	Manufacture of rubber and plastic products;
26	Manufacture of other non-metallic mineral products;
27	Manufacture of basic metals;
	MACHINERY
28	Manufacture of fabricated metal products, except machinery and equipment;
29	Manufacture of machinery and equipment n.e.c.
30	Manufacture of office machinery and computers;
31	Manufacture of electrical machinery and apparatus n.e.c.
32	Manufacture of radio, television and communication equipment;
33	Manufacture of medical, precision and optical instruments, watches and clocks;
34	Manufacture of motor vehicles, trailers and semi-trailers;
35	Manufacture of other transport equipment;
36	Manufacture of furniture; manufacturing n.e.c.
	SERVICES RECREATIONAL
40	Electricity, gas, steam and hot water supply;
72	Computer and related activities;
74	Other business activities;
92	Recreational, cultural and sporting activities;
93	Other service activities.

Appendix 5: Mores stylized facts at a country level.

Geographic and product extensive margins (sorted by descending GDP)



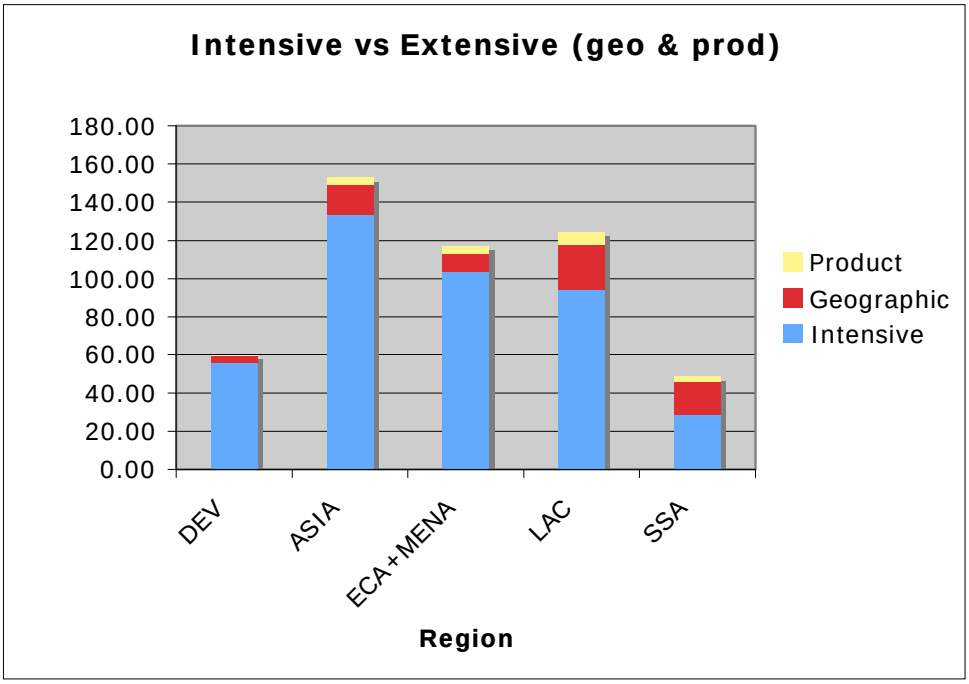


Table 6: Decomposition of the extensive margin of adjustment (year base=1995)

	EXTENSIVE (2)	TOTAL OPND (2.1)	TOTAL NP (2.2)	NPOD (2.2.1)	NPND (2.2.2)
ARG	39.03	**30.73**	**8.30**	0.32	7.98
AUS	15.32	**14.69**	**0.63**	0.09	0.53
BGD	8.85	**5.80**	**3.06**	0.80	2.25
BRA	24.86	**14.87**	**9.99**	0.21	9.78
CAN	2.12	**1.95**	**0.17**	0.19	-0.02
CHL	17.05	**14.20**	**2.86**	0.16	2.70
CHN	14.08	**12.90**	**1.17**	0.23	0.94
EEC15	1.47	**1.50**	**-0.04**	-0.02	-0.02
GHA	11.33	**10.17**	**1.16**	0.02	1.13
IDN	21.38	**15.34**	**6.04**	2.50	3.54
IND	25.61	**20.27**	**5.34**	2.46	2.88
JPN	0.60	**0.56**	**0.04**	0.00	0.04
KEN	8.14	**3.41**	**4.73**	0.43	4.30
MAR	9.52	**8.05**	**1.46**	0.04	1.42
MEX	4.59	**4.01**	**0.58**	0.25	0.33
MUS	3.13	**2.07**	**1.06**	-1.99	3.04
MYS	11.27	**9.00**	**2.27**	1.10	1.17
PER	33.06	**19.92**	**13.14**	1.20	11.94
THA	20.73	**15.61**	**5.13**	1.29	3.84
TUN	8.85	**3.75**	**5.10**	2.04	3.06
TUR	14.38	**9.95**	**4.44**	2.59	1.85
UGA	10.28	**6.02**	**4.26**	0.01	4.25
USA	0.46	**0.46**	**0.00**	0.00	0.00
ZAF	51.72	**49.45**	**2.27**	1.11	1.16
		11.45	**3.46**	**0.63**	**2.84**

Table 7: Decomposition of export growth (benchmark year=1995)

	Intensive trade (1)	Extensive trade (2)	Overall Change (3)	Extensive as % Total (2) / (3)
ARG	32.44	39.03	71.46	45.39%
AUS	64.62	15.32	79.93	19.16%
BGD	152.96	8.85	161.81	5.47%
BRA	89.65	24.86	114.52	21.71%
CAN	61.98	2.12	64.10	3.31%
CHL	98.37	17.05	115.43	14.77%
CHN	302.40	14.08	316.48	4.45%
EEC15	77.80	1.47	79.27	1.85%
GHA	14.00	11.33	25.33	44.73%
IDN	57.30	21.38	78.68	27.18%
IND	113.40	25.61	139.01	18.42%
JPN	31.83	0.60	32.43	1.86%
KEN	34.09	8.14	42.23	19.27%
MAR	43.45	9.52	52.97	17.97%
MEX	142.85	4.59	147.44	3.11%
MUS	12.84	3.13	15.98	19.61%
MYS	91.41	11.27	102.68	10.98%
PER	108.41	33.06	141.47	23.37%
THA	83.77	20.73	104.51	19.84%
TUN	78.86	8.85	87.71	10.09%
TUR	189.76	14.38	204.14	7.05%
UGA	-40.41	10.28	-30.14	-34.10%
USA	41.77	0.46	42.23	1.09%
ZAF	123.51	51.72	175.23	29.52%
	83.63	**14.91**	**98.54**	**0.14**

Table 8: Breakdown Extensive Margin of adjustment

	EXTENSIVE	OPND (3)	NP (4)	NPOD (5)	NPND (6)
LAC	**23.72**	**16.75**	**6.97**	**0.43**	**6.54**
DEV	**3.99**	**3.83**	**0.16**	**0.05**	**0.11**
SSA	**16.92**	**14.22**	**2.69**	**-0.08**	**2.78**
ASIA	**16.99**	**13.15**	**3.84**	**1.40**	**2.44**
ECA+MENA	**10.92**	**7.25**	**3.67**	**1.56**	**2.11**